14/12/22   (170)

K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)
Subject Code: HAIMLC501 /HDSC501
Subject Name: Mathematics for AI & ML/Mathematics for Data Science   Date:

<table>
<tr><td colspan="2" align="center">Nov – Dec 2022
Program: B. Tech
Examination: TY   Semester: V
Course Code: HAIMLC501 /HDSC501 and
Course Name: Mathematics for AI & ML/Mathematics for Data Science
Duration: 2.5 Hours</td><td colspan="3" align="center">Max. Marks: 60</td></tr>
<tr><td colspan="5">Instructions:
(1) All questions are compulsory.
(2) Draw neat diagrams wherever applicable.
(3) Assume suitable data, if necessary.</td></tr>
<tr><td></td><td></td><td>Max. Marks</td><td>CO</td><td>BT level</td></tr>
<tr><td>Q 1</td><td>Solve any six questions out of eight:</td><td>12</td><td></td><td></td></tr>
<tr><td>i)</td><td>If $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, find the eigen values of A and eigen vectors for $\lambda = 5$ only.</td><td>2</td><td>1</td><td>3</td></tr>
<tr><td>ii)</td><td>State central limit theorem.</td><td>2</td><td>2</td><td>3</td></tr>
<tr><td>iii)</td><td>State the type of variables in each of the following exercises. Are they discrete or continuous?
a) Weight of newspapers recycled on a single day
b) Number of claims filed with an insurance company during a single day
c) Time required to complete a questionnaire
d) Number of consumers in a poll of 1,000 who consider nutritional labeling on food products to be important</td><td>2</td><td>3</td><td>3</td></tr>
<tr><td>iv)</td><td>State four benefits of feature engineering.</td><td>2</td><td>4</td><td>3</td></tr>
<tr><td>v)</td><td>Find the optimum value of x for the function $f(x) = 12x - 3x^4 - 2x^6$ with initial value $x_i = 1$ for first iteration only using Newton's Method.</td><td>2</td><td>5</td><td>3</td></tr>
<tr><td>vi)</td><td>Which are the reasons for poor data quality?</td><td>2</td><td>4</td><td>3</td></tr>
<tr><td>vii)</td><td>Parking on Campus Six vehicles selected from a campus vehicle database are shown in the table.</td><td>2</td><td>3</td><td>3</td></tr>
</table>

K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Subject Code: HAIMLC501 /HDSC501

Subject Name: Mathematics for AI & ML/Mathematics for Data Science   Date:

<table>
<tr><td colspan="3"></td></tr>
</table>

						One-way Commute Distance (kilometers)	Age of Vehicle (years)			
	Vehicle	Type	Make	Carpool?						
	1	Car	Honda	No		37.8	6			
	2	Car	Toyota	No		27.5	3			
	3	Truck	Toyota	No		16.2	4			
	4	Van	Dodge	Yes		50.7	2			
	5	Motor-cycle	Harley-Davidson	No		40.8	1			
	6	Car	Chevrolet	No		8.6	9			

a. List the variables that are being measured. What types are they?
b. Is this univariate, bivariate, or multivariate data?

viii)	What is Dimensionality Reduction?	2	6	3
Q.2	Solve any four questions out of six.	16		
i)	Apply Gauss Elimination Method to solve $x + 3y - 2z = 5$, $2x + y - 3z = 1$, $3x + 2y - z = 6$.	4	1	3
ii)	The means of two independent samples of size 8 and 7 are 1134 and 1024 respectively. The standard deviation of these two samples is 35 and 40 respectively. What is the value of test statistic t in order to test the significance of difference between sample means?	4	2	3
iii)	The Dow Jones Industrial Average was monitored at the close of trading for 10 days in a recent year with the results shown in following table. Dow Jones Industrial Average Draw the line chart.	4	3	3
iv)	Find the optimum value of of $f(x) = x^4 - 14x^3 + 60x^2 - 70x$ in the interval [0,2] upto 4th iteratin using bisection method.	4	5	3
v)	Explain various methods for handling missing values.	4	4	3
vi)	Explain when should we use PCA?	4	6	3
Q.3	Solve any two questions out of three.	16		

Dow Jones Industrial Average (table in iii):

Day	1	2	3	4	5	6	7	8	9	10
DJIA	21,479	21,478	21,320	21,414	21,409	21,532	21,553	21,638	21,630	21,575

K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Subject Code: HAIMLC501 /HDSC501

Subject Name: Mathematics for AI & ML/Mathematics for Data Science   Date:

i)	Find a root of an equation $f(x) = x^3 - x - 1$ using Regula Falsi method.	8	5	3

ii)	A certain drug administered to 12 patients resulted in the following change in them
	    Blood Pressure 5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4
	    Can we conclude that drug increase the Blood Pressure?
		8	2	3

iii)	The qualitative variable "class status" has been recorded for each of 105 students in an introductory statistics class, and the frequencies are shown in following table 	Status	Freshman	Sophomore	Junior	Senior	Grad Student
---	---	---	---	---	---
Frequency	5	23	32	35	10 | Draw Pie and Bar Charts for status.	8	3	3
Q.4	Solve any two questions out of three.	16		
i)	Find the dimension and basis for the four fundamental subspaces for $A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}_{3\times 4}$	8	1	3
ii)	Explain the processes involved in feature engineering.	8	4	3
iii)	Explain how the Linear Discriminant Analysis (LDA) work?	8	6	3

****